

2/5/1 (Item 1 from file: 347)
DIALOG(R) File 347: JAPIO
(c) 2000 JPO & JAPIO. All rts. reserv.

02350976 **Image available**
PICTURE REGISTER SYSTEM

PUB. NO.: 62-267876 A]
PUBLISHED: November 20, 1987 (19871120)
INVENTOR(s): MURAKAMI TATSUYA
NAKANO YASUAKI
FUJINAWA MASAOKI
FUJISAWA HIROMICHI
APPLICANT(s): HITACHI LTD [000510] (A Japanese Company or Corporation), JP
(Japan)
APPL. NO.: 61-110769 [JP 86110769]
FILED: May 16, 1986 (19860516)
INTL CLASS: [4] G06F-015/62
JAPIO CLASS: 45.4 (INFORMATION PROCESSING -- Computer Applications)
JOURNAL: Section: P, Section No. 699, Vol. 12, No. 151, Pg. 51, May
11, 1988 (19880511)

ABSTRACT

PURPOSE: To perform the input jobs against two picture data bases independently of each other and collectively and to improve the picture registering efficiency, by securing automatically the correspondence between the pictures on those data bases and the register items on a register item data base.

CONSTITUTION: A recognizing part 140 recognizes F as the contents of an entry column A 212 and then recognizes characters as the contents of both columns B 213 and C 214 to obtain the sorting classes and document numbers. Thus these classes and numbers are allocated to all pictures inputted until the input of a trailer sheet 250. While the register item data stored in a register item data base 190 and having the sorting classes and document numbers equal to those allocated to the pictures are set opposite to all pictures belonging to a relevant single document. When the front side of the document is read by a normal automatic paper feeder, the rear sides are successively read as 550, 540, ... in many cases. Therefore the rearrangement of documents is not needed in an input mode if the page numbers are produced in the order of 2N-1, 2N-3, ..., 3, 1.

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A)

昭62-267876

⑬ Int. Cl.⁴

識別記号

庁内整理番号

⑭ 公開 昭和62年(1987)11月20日

G 06 F 15/62

6615-5B

審査請求 未請求 発明の数 1 (全6頁)

⑮ 発明の名称 画像登録方式

⑯ 特 願 昭61-110769

⑰ 出 願 昭61(1986)5月16日

⑱ 発 明 者 村 上 達 也 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内
 ⑲ 発 明 者 中 野 康 明 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内
 ⑳ 発 明 者 藤 縄 雅 章 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内
 ㉑ 発 明 者 藤 沢 浩 道 国分寺市東恋ヶ窪1丁目280番地 株式会社日立製作所中央研究所内
 ㉒ 出 願 人 株式会社日立製作所 東京都千代田区神田駿河台4丁目6番地
 ㉓ 代 理 人 弁理士 小川 勝男 外1名

明 細 書

1. 発明の名称

画像登録方式

2. 特許請求の範囲

1. 文書画像情報を登録し、登録済の画像ファイルを検索する画像ファイル装置において、キーワード等検索に用いるデータとその検索用データと登録された文書一件ごとの画像を対応づけるための文書認識データとを含む登録データと、該データを蓄積する登録事項データベースと、上記画像イメージを蓄積する画像データベースと、上記登録する画像とその画像に関する登録データとを対応づけるための文書認識データを画像データに付加する手段と、該付加された文書認識データを認識する認識部と、該認識された文書認識データと上記登録データ内の文書認識データとを対応づける手段と、該対応づけられた結果より画像データとその画像に関する登録データとを文書一件ごとに対応づける手段と、一枚もしくは複数の画像より構成される一件の

文書内で個々の画像に対してそれぞれ別の頁番号を自動生成し割り当てる手段と、上記登録データによって指定された文書からさらに頁番号により個々の画像を指定する手段を有することを特徴とする画像登録方式。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は文書等の画像をファイルに蓄積する画像ファイル装置への画像登録方式に係り、特に一度に大量の画像を効率よく登録する方式に関する。

〔従来の技術〕

画像ファイル装置に文書等の画像を登録する場合、検索のためのキーワードやUDCコード等をデータ・ベースとして登録することが不可欠である。また効率のよい検索を行うためには原文の題名や著者名等の書誌的事項もキーワード等と共に登録する必要がある。

従来の登録においては、画像を一枚入力する枚に、あるいは一度に複数枚の画像のみを入力した後それらの画像を一枚ずつ表示しながら、キーワ

ードや書誌的事項（これらを総称して以後、登録事項と記す）を人手により入力していた（例えば、特開昭56-22225号公報参照）。また一件の文書が多数頁の画像より構成されている合本形式の文書を登録する場合は登録事項の入力は文書1件あたり1度でよい。しかし後に原文書画像を表示する際本来の順序で（正しい頁順で）表示するためには、画像の入力を頁順序に従って行うか、画像一枚毎に頁番号を入力する必要がある。

一方ファイルに入力された画像にそれぞれ入力順に一連の番号（文書ID番号）を割り当て、その文書ID番号を検索コードとして登録し、人手による登録事項の入力を不要にする方式が実用化されている。しかしながらこの方式は同一ファイル上に複数種類の文書を登録する場合等については配慮されていない。

また入力する画像上の文字列の特定領域を指定し、その特定領域の文字列を認識し、得られた結果を検索コードとして登録する方式も種々知られている（例えば、特開昭60-17565号公報

りに行うか、全ての画像に対して逐一頁番号を入力する必要がある。しかし合本形式の文書の多くは紙等記録媒体の表裏両面に記録されており、これを頁番号順に入力するには1枚の記録媒体の表と裏を続けて入力しなければならないため画像入力の自動化も困難で効率が悪い。また全ての画像について頁番号を逐一入力する方式は画像を一枚ずつ全て表示確認する必要があり効率が悪く、頁数が多い時はミスも生じやすい。

一連の番号を画像の入力順に割り当て、検索コードとしてそれだけを自動的に登録する方式は登録の効率は高い。しかしこの方式では検索に用いることのできる情報は入力した順序だけであり、内容による検索等はできない。

入力した画像上の特定領域の文字列を検索コードとして指定し、その文字列を認識することで登録事項を自動的に入力する方式は単葉形式の文書を登録する場合には効率がよい。

しかしこの方式は合本形式の文書登録については配慮していないため前記した頁順の問題がやは

参照）。しかしこの方式は一件の文書が一枚の画像からなる単葉形式の文書を登録するためのものであり、合本形式の文書を登録する場合等一件の文書に属する複数の画像に同一の登録事項を対応させることについては配慮されていなかった。

（発明が解決しようとする問題点）

上記従来技術において登録事項の入力を文書一件ごとに行う方式は画像の入力とキーボード等からの登録事項の入力を交互に行うため登録の効率が悪かった。また一度入力した画像を表示しながら登録事項を入力する方式も表示・確認・入力という処理を交互に行うため効率は悪い。さらに画像と登録事項を正しく対応させるには、登録すべき内容を熟知した人間が登録作業を行う必要があった。

合本形式の文書を登録した場合、表示の際には複数枚の画像を次々と表示することが求められるが、この時表示される画像の順序は登録したものと文書の順序と同じにする必要がある。従来そのためには、画像の入力を登録する文書の頁番号通

り存在し、また同一文書内の各頁に対して検索コードの登録をう場合等問題がある。

本発明の目的は入力された画像と登録事項とを自動的に対応させることにより、画像入力と登録事項入力をそれぞれまとめて行えるようにし、さらに登録する文書が合本形式の場合に入力しやすい順序で登録しても、各画像に対して自動的に正しい頁番号を割り当てることにより、画像の自動入力を可能とする効率のよい登録方式を提供することにある。

（問題点を解決するための手段）

上記の目的を達成するために、以下の手段を用いる。まず画像検索用の登録事項のデータ・ベースに、個々の文書に対しその文書の内容によって登録者等が分類した分類クラスと、各分類クラスの中で個々の文書を特定する文書番号も記入する。また画像の入力に際しては、はじめに入力する文書の分類番号や文書番号等を記入した登録用記録媒体（以後ヘッダーシートと記す）を入力する。次にヘッダーシート上の記載事項に文字認識を行

い、その結果と登録事項データベースを対応させる。またヘッダーシートには入力する文書が単票形式か合本形式かを示す情報と、合本形式の場合には画像が両面に記載されているか片面にだけかという記載方式の違いを示す情報も記入しておく。そして入力文書が合本形式の場合にはその記載方式の文書にとって最も入力しやすい順序で入力した時に応じた順序の頁番号を自動生成し各画像に対応させる。

〔作用〕

登録事項データベースには対応する画像の検索コード及び書誌的事項の他そのデータと画像を結びつけるために分類クラスや文書番号等の情報も登録する。一方向画像入力の際には登録すべき文書の入力に先だって分類クラスや文書番号を記入したヘッダーシートを入力する。次にこのヘッダーシートの内容を既知の文字認識の手法により認識して分類クラスや文書番号を入力する。この情報と同じものを登録事項データベースの中より選り出すことにより画像と登録事項データベースを自

動的に対応づけることができる。これにより登録に必要な画像入力と登録事項の入力を別々にそれぞれまとめて行えるので、効率よく登録できる。また画像と登録事項の対応づけは自動的に行うので、登録作業を行う人が内容について熟知している必要はない。

また、ヘッダーシートに文書の形式（単票か合本か）や記載方式（片面文書か両面文書か）を記入してそれらの情報も文字認識により入力する。その結果もし合本形式であるならば、1件の文書中の各画像にはそれぞれ頁番号をつける必要がある。合本形式の文書は入力しやすい順序で入力した時、記載方式の違いにより入力された頁の順序が異なる。そこでヘッダーシート上の情報よりその文書の記載方式を入力し、それに応じた順序で頁番号を生成することにより、文書の記載方式を問わず、通常の自動給紙装置で入力できるようにして入力の効率を高める。

〔実施例〕

以下に図を用いて本発明の詳細な説明を行う。

第1図は本発明の一実施例を示すブロック図である。図中110は周知の機構により複数枚の原文書画像を連続的に読み取ることでできる機能を持った読取部、120は読み取った画像の信号にノイズ除却等の処理を加える画像処理部、130は画像1枚分以上の内蔵を有する画像メモリ、140は画像メモリ130中の信号より必要な情報を抽出し内容を認識する認識部、150は装置全体の動作や処理を制御する制御部、160は画像メモリ上の画像を表示する出力部、170は装置に指示やデータを入力するキーボード等の入力部、180は読取部110より入力された画像を蓄積する画像データベース、190は入力部170等より入力された登録事項を蓄積する登録事項データベースで、200は各部分をつなぐバスラインである。登録事項データベース190には個々の文書に対応する登録データが文書の件数分蓄積されている。登録データは検索に用いる登録事項と、画像データベース上のデータとの対応に用いる接続データより構成されている。さらに

登録事項は文書の題名や著者名等の書誌的事項と、検索用のキーワードより構成され、一方接続データは分類クラスと文書番号より構成されている。

第2図は原文書画像の登録手順を示す。この場合登録する文書は合本形式の文書を例にしている。210はヘッダーシート、220、230、240はそれぞれ登録すべき合本文書の一部である原文書画像で、250はヘッダーシートと同じ書式で文書入力の終了を示すシート（以後トレーラ・シートと記す）である。入力作業は左から順にまずヘッダーシートを入力し、続いて原文書画像を文書1件分まとめて入力し、最後にトレーラ・シートを入力して終了する。

ヘッダーシート上には入力する文書に関する情報の記入枠211が記されている。第3図に記入枠211の拡大図を示す。記入欄A212には登録する文書の形式及び記載方式を表す記号を記入し、記入欄B213には分類クラス、記入欄C214には文書番号を記入する。記入欄A212に記入する記号の一例を第4図に示す。

第4図に示す通りこの実施例では文書の形式や記載方式の違う3種類の文書に対して登録を行う。まず初めに合本形式の両面文書を登録する場合の処理について説明する。この方式は1件の文書を片面ずつまとめて入力するもので、この場合情報記入欄A212に記入する記号は第4図に示す通りFとBの2種類あるが、それぞれ文書の表側を入力する時と裏側を入力する時に用いるものである。両者の違いは生成する頁番号が奇数であるか偶数であるかだけである。仮に原文書の表側から入力したとすると、記入欄A212にはFを記入する。認識部140が記入欄A212の内容としてFを認識すると、続いて記入欄B213および記入欄C214の内容に文字認識を行い、分類クラスと文書番号を得る。これにより以後トレーラシート250が入力されるまでに入力された全ての画像にはこの分類クラスと文書番号が割り当てられる。また登録事項データベース180中でこれと同じ分類クラスと文書番号を有する登録データがこの1件の文書に属する全ての画像に対応づけられる。

側のみを連続して入力したのであるから、N枚の原文書画像が入力されたとすると、割り当てられる頁番号は1, 3, 5, ..., $2N-3$, $2N-1$ と1頁おきの値となる。また一般に文書の多くは片面文書・両面文書を問わず第5図に示すよう文面(両面文書の場合は表側)を手前に向けた場合、頁番号の少い順に500, 510, ..., 550と手前から並べられている事が多い。一方通常の自動給紙装置で第5図の文書の手前側の面を読む場合、550, 540, ...と後方より読み取ることが多い。したがって頁番号は $2N-1$, $2N-3$, ..., 3, 1の順に生成すれば入力時に文書を並びかえる必要がなくなる。こうして生成された頁番号は登録事項データベース中のこの文書に対応する登録データに順番に入力される。

続いて裏面の文書を入力する。この場合登録上の処理は表側と同じであり、相違点はヘッダーシート210の記入欄212にBと記入すること、生成される頁番号のみである。表側を自動給紙装置により入力した場合、文書の並び順は入力前と

認識終了後制御部150の指示により2枚目の画像が読取られ、ヘッダーシート210と同様の処理を経て画像メモリ130に記録される。この画像に対して認識部140は予め指定された場所に記入枠251あるかどうかのみをチェックする。枠がないと判断された場合その画像は画像データベース上に蓄積され、その蓄積場所が登録事項データベースに登録される。同様にして1件の文書の表側だけ全て入力し終ると、トレーラシート250が入力される。トレーラシート250では認識部140により枠251が見つけれられる。すると次に認識部140は記入欄D252に文字認識を行い、ここで入力終了を示す記号が認識されると入力は終る。

しかしながらこの時点では登録事項データベース中の登録データは1件の文書全体に対して対応しているだけなので、個々の画像を指定することはできない。蓄積された各画像を個別に指定するには画像に頁番号を割り当て、それも登録する必要がある。この場合は両面に記載された文書の表

逆になる事が多い。したがって生成する頁番号は2, 4, ..., $2N-2$, $2N$ とする。こうして合本形式の両面記載文書の登録が終了する。

合本形式の片面記載文書を登録する場合の処理もほぼ同様であり、記入欄212の認識により記号Gが確認されると生成される頁番号がN, N-1, ..., 2, 1となる。

次に単票形式の文書を登録する場合について説明する。この場合一枚の文書画像から成る文書を複数件連続して入力することになる。したがって各文書ごとに文書番号が異なるが、頁番号は生成する必要がない。

第6図は単票形式の登録手順を示す図である。ここでは例として分類クラスの同じ複数の文書を登録する場合について説明する。入力は単票形式の文書を登録することを示す記号SGを記入欄A212に記したヘッダーシート210より行われる。認識部140は記入欄A212の内容として記号SGを認識すると次に記入欄B213の内容に対し文字認識を行い、分類クラスを認識する。

以後トレーラシート250が入力されるまでに入力された全ての文書にはこの分類クラスが割り当てられる。続いて順次文書入力し画像データベースに登録するが、この時文書と同時に文書番号も入力する必要がある。そこで原票形式文書に登録する場合、ヘッダシートと同じフォーマットの入力シート620の記入欄C214に文書番号のみを記入したものの上に登録する原票の原文書画像630を添付し入力する。認識部140は2枚目以降の文書に対してまず記入欄A212を調べ、トレーラシートか否かを判定し、文書の場合は記入欄C214に文字認識を行い、文書番号を認識する。この文書番号と分類クラスから合本形式の場合と同様に、登録事項データベース180上の登録データと画像を対応づけることができる。

このようにして登録すべき文書の形式、記載方式を問わず、効率よく登録できる。

〔発明の効果〕

以上詳細に説明した通り、本発明によれば、画像データベース上の画像と、登録事項データベ

ス上の登録事項を自動的に対応させられるので2つのデータベースに対する入力作業は別々にまとめて行うことができる。また合本形式の文書に登録する場合にも自動給紙装置による連続入力が可能になる。これらの結果画像登録作業の効率を高めることができる。

4. 図面の簡単な説明

第1図は本発明の一実施例を示すブロック図、第2図は合本形式の文書に登録する際の入力順序を示す図、第3図はヘッダシート上の原文書画像に関する事項の記入欄を示す図、第4図は原文書の形式及び記載方式と各場合における形式記号の例を示す図、第5図は自動給紙装置による画像読取りと入力される頁番号の順序を示す図、第6図は原票形式の文書に登録する際の入力順序を示す図である。

110…読取部、130…画像メモリ、

140…認識部、150…制御部、180…画像

データベース、190…登録事項データベース、

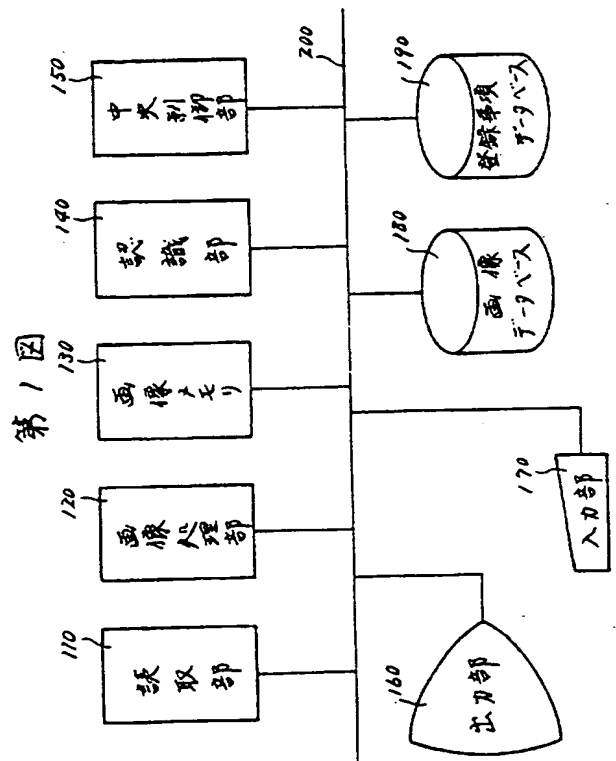
210…ヘッダシート、220…原文書画像、

250…トレーラシート、211…記入枠、

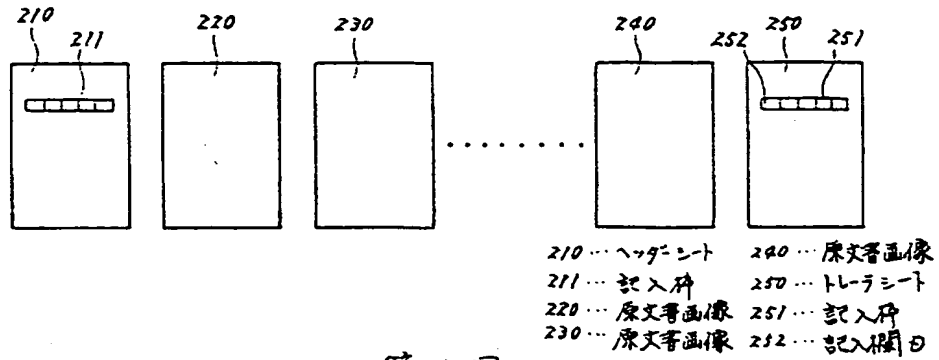
212…記入欄A、620…入力シート、

630…原文書画像。

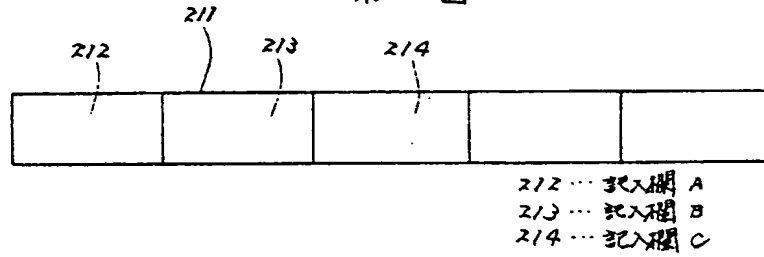
代理人 井理士 小川勝男



第 2 図



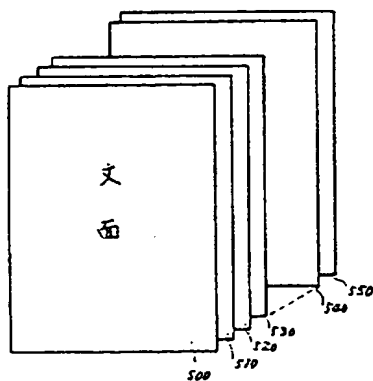
第 3 図



第 4 図

形式	記載法	表	記号
合本	両面	表	F
		裏	B
単葉	片面	/	G
		/	SG

第 5 図



第 6 図

